2-О

УДК 681.518.25

Ю.К. Орлов, О.М. Кир'ян
Державний університет інформатики і штучного інтелекту, м. Донецьк, Україна

Прогнозування залишків коштів на поточних і депозитних рахунках клієнтів

У статті розглядаються проблеми побудови моделі прогнозування залишків коштів на поточних і депозитних рахунках клієнтів. Пропонується побудова системи прогнозування за допомогою нейронних мереж, що дозволяє підвищити якість прогнозу.

Вступ

Ліквідність – один з основних і найбільш складних факторів, що визначають фінансовий стан і надійність банку. Величезний вплив на банківську ліквідність має така категорія пасивів, як депозити й поточні рахунки клієнтів, про що свідчить занадто висока питома вага таких пасивів у балансах українських комерційних банків. Залишки на поточних та депозитних клієнтських рахунках є найдешевшим, а тому основним, джерелом коштів, використовуваних банками для здійснення кредитної політики, чим сприяють максимізації банківського прибутку шляхом значного зменшення процентних витрат банку. Однак дана категорія банківських ресурсів характеризується високою рухливістю коштів і пов'язана з ризиком ліквідності більше, ніж будь-яка інша категорія, тому що строки вилучення коштів з розглянутих рахунків є невизначеними за своєю природою, отже, визначення очікуваного залишку коштів на рахунках допоможе банкам якомога раціональніше використовувати це досить прибуткове джерело залучених коштів, мінімізуючи при цьому ризик ліквідності банку, а також встановити напрями депозитної політики банку [1].

Таким чином, проблема прогнозування передбачуваного залишку коштів на поточних і депозитних рахунках займає особливе місце в процесі керування банківською ліквідністю, дає можливість, незважаючи на високу рухливість коштів на таких рахунках, визначити постійний гарантований їхній залишок і використати його як стабільний кредитний ресурс [2].

Мета даної роботи полягає в розробці моделі прогнозування коштів на поточних і депозитних рахунках клієнтів з використанням апарату нейронних мереж (НМ).

Постановка задачі побудови моделі прогнозування

Спостерігаються величини $Y = Y(t)$ середньозважених залишків коштів на рахунках клієнтів у дискретні моменти часу (місяці) $t_i = t_0+i\Delta t$, $i = \overline{1,T}$, де t_0 – деякий початковий момент часу, Δt – проміжок часу між спостереженнями, T – горизонт прогнозування. Задача прогнозування полягає в тому, щоб за відомою ділянкою часового ряду $\left\{ t_i,\ Y(t_i)\right\}_{i=1}^{T}$ оцінити майбутні значення величини Y на заданий період попередження прогнозу, з огляду на набір факторів $X = (X_1(t), X_2(t), \ldots X_n(t))$, що впливають на прогнозовану величину.

Для побудови моделі поведінки залишків коштів на депозитних рахунках необхідно відповідним чином описати її. Для збору відомостей про досліджуваний процес у першу чергу важливе розуміння сутності самого процесу. Очевидно, що ринок депозитів, а отже, й попит на них існує й розвивається не сам по собі, а під впливом різних факторів, як внутрішніх, так і зовнішніх. Залишок на рахунках, що аналізуються, у зв'язку з високою мобільністю коштів не постійний і часто вкрай мінливий. Числове значення залишку на всіх балансових рахунках даної групи являє собою середньозважену суму залишків на деякій кількості рахунків у відповідний момент часу і є випадковою величиною, безперервною в інтервалі $(0;+\infty)$[3]. Згідно із законом великих чисел при деяких досить широких умовах (попарна незалежність випадкових величин, рівність їхнього математичного сподівання, рівномірна обмеженість дисперсій) сумарне поводження значної кількості випадкових величин практично стає закономірністю.

Керуючись цими міркуваннями, з метою збільшення якості вихідних даних для прогнозування, а отже, і поліпшення якості самого прогнозу необхідно зробити класифікацію клієнтів комерційного банку залежно від параметрів часового ряду, що описує поводження залишків на кожному конкретному рахунку шляхом розбивки клієнтів на групи за величиною їхнього математичного сподівання, що характеризує середню суму коштів, що зберігаються на рахунку в аналізований період часу. Потім кожна група розбивається на підгрупи залежно від величини дисперсії. Прогнозування передбачається здійснювати по кожній групі окремо, представляючи загальний результат сумою отриманих прогнозів.

Вибір моделі прогнозування був обумовлений перевагами апарату нейронних мереж порівняно з іншими існуючими методами. При цьому на підвищення ефективності навчання нейронної мережі істотний вплив робить передобробка даних, основним принципом якої є збільшення інформативності прикладів.

Для розв'язання поставленого завдання прогнозування методами кореляційного аналізу визначається набір факторів, що впливають на прогнозований показник. Потім вхідний простір ознак повинен бути оптимізований на підставі методики box-counting шляхом визначення індивідуальної значимості (крос-ентропії) для кожного входу. При необхідності відновлюється пропущена інформація (шляхом заміни пропущених значень на середнє значення величини), усуваються аномальні викиди, які можуть спотворити результат навчання (шляхом перетворення викидів у пропущені значення). Далі провадиться нормування вхідних даних в інтервалі [–1;1], а для досягнення максимальної інформаційної насиченості кожного із вхідних факторів окремо і їхньої статистичної незалежності повинна бути здійснена декореляція («вибілювання») входів [4].

Первинний відбір факторів, що впливають, здійснювався на підставі простих теоретичних міркувань, при цьому враховувалася практична можливість одержання даної інформації, а також передбачалася можливість її наступного доповнення. Як вхідні фактори запропоновані наступні величини:

X_1 – рівень інфляції (індекс споживчих цін на товари й послуги), %;
X_2 – валовий внутрішній продукт, млн грн.;
X_3 – середньомісячна номінальна заробітна плата працівників, грн.;
X_4 – індекс реальної заробітної плати працівників, %;
X_5 – доходи населення, млн грн.;
X_6 – загальний рівень безробіття (на кінець періоду), %;
X_7 – курс долара США, грн. за 100 од. валюти;
X_8 – курс ЄВРО, грн. за 100 од. валюти;
X_9 – кількість відкритих рахунків у даному місяці;

X_{10} – середньозважена процентна ставка по даному типу депозитів, %;

X_{11} – ставка рефінансування НБУ, %;

X_{12} – темпи росту грошової маси, млрд грн.

Нехай m – ширина вхідного вікна, p – вихідного, s = 1 – величина кроку, N – кількість спостережень. Якщо припустити, що $X_{K+1}(t_i) = Y(t_i)$, $X_{K+2}(t_i) = \dfrac{t_i}{T}$, $i = \overline{1,T}$, де $T = N–1–p$, K – кількість врахованих факторів, та положити $K = K+2$, то множина даних для розв'язання задачі прогнозування буде мати вигляд, що наведений у табл. 1.

Таблиця 1 – Множина даних для задачі прогнозування

Входи										Виходи		
Фактор X_1			Фактор X_2			..	Фактор X_K			Y		
$X_1(t_1)$		$X_1(t_m)$	$X_2(t_1)$		$X_2(t_m)$	..	$X_K(t_1)$		$X_K(t_m)$	$Y(t_{m+1})$		$Y(t_{m+p})$
$X_1(t_2)$		$X_1(t_{m+1})$	$X_2(t_2)$		$X_2(t_{m+1})$		$X_K(t_2)$		$X_K(t_{m+1})$	$Y(t_{m+2})$		$Y(t_{m+p+1})$
…		…	…		…		…		…	…		…
$X_1(t_{T-m+1})$		$X_1(t_T)$	$X_2(t_{T-m+1})$		$X_2(t_T)$	..	$X_K(t_{T-m+1})$		$X_K(t_T)$	$Y(t_{T+1})$		$Y(t_{T+1+p})$

Побудова системи прогнозування

Система прогнозування складається з декількох нейромережних модулів, кожний з яких є багатошаровим персептроном і використовується для дослідження рахунків певної групи клієнтів, тобто навчання співвідношень між перерахованими вище економічними факторами й величинами середньозважених щомісячних залишків на рахунках клієнтів цієї групи. Таким чином, входами кожного модуля є виявлені на підставі вищеописаних критеріїв і проведених експериментів фактори, відомості про залишки на певній групі рахунків за n попередніх місяців (а саме – їхня відносна зміна стосовно попереднього місяця), а також відомості про сезонність, тобто дані про середньозважені щомісячні залишки відповідного періоду попереднього року. Виходом є прогнозоване значення залишку коштів у планованому місяці, одержуване як сума прогнозів по кожній групі (рис. 1).

Розв'язання задачі виконується відповідно до схеми, представленої на рис. 2.

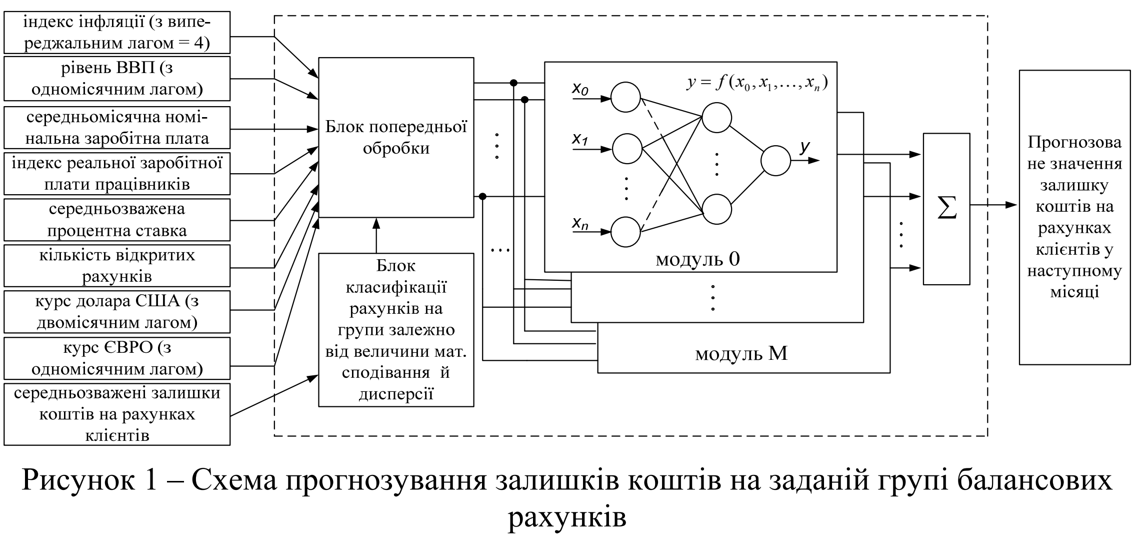


Рисунок 1 – Схема прогнозування залишків коштів на заданій групі балансових рахунків

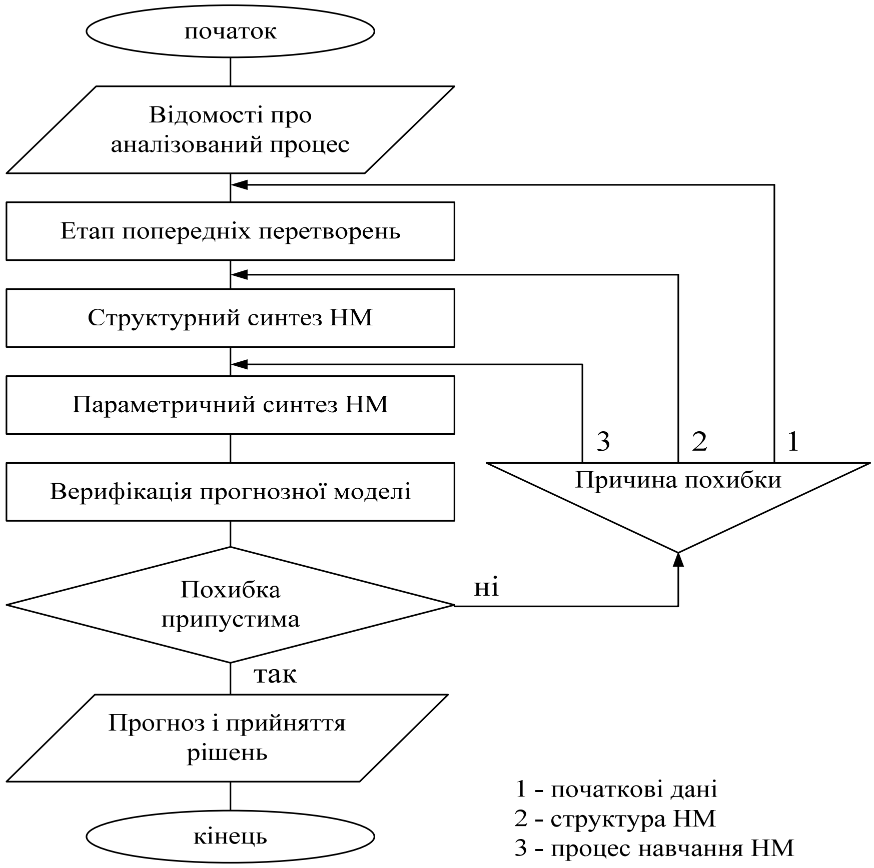


Рисунок 2 – Укрупнена схема розв'язання задачі прогнозування

Для навчання мережі використовується алгоритм зворотного поширення помилки, що будується таким чином:

1. Ініціювати вагові коефіцієнти мережі.

2. Подати на входи мережі один з можливих образів і розрахувати значення виходів:

$$y_j^{(n)} = f\left(\sum_{i=0}^{M} y_i^{(n-1)} \cdot w_{ij}^{(n)}\right), \tag{1}$$

де М – число нейронів у шарі $n–1$; $y_i^{(n-1)}=x_{ij}^{(n)}$ – i-й вхід нейрона j шару n, $w_{ij}^{(n)}$ – вага від нейрона i шару $n – 1$ до нейрона j шару n, $f(s)$ – сигмоїд.

3. Розрахувати нев'язання нейронів для корекції вагових коефіцієнтів по формулах:

$$\delta_j^{(n)} = \left[\sum_k \delta_k^{(n+1)} \cdot w_{jk}^{(n+1)}\right] \cdot \frac{dy_j}{ds_j} \quad \text{– для схованих шарів;} \tag{2}$$

$$\delta_l^{(N)} = (y_l^{(N)} - d_l) \cdot \frac{dy_l}{ds_l} \quad \text{– для вихідного шару.} \tag{3}$$

4. Розрахувати похідну помилки за вагою між i-м нейроном шару $n – 1$ й j-м нейроном шару n:

$$\frac{\partial E}{\partial w_{ij}} = \delta_j^{(n)} \cdot y_i^{(n-1)}. \tag{4}$$

5. Скорегувати всі ваги нейронної мережі: $w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + \Delta w_{ij}^{(n)}(t)$, де зміни ваг $\Delta w^{(n)}$ шару n визначити відповідно до обраного методу корекції вагових коефіцієнтів.

6. Якщо подані всі вхідні образи, перейти до наступного кроку, інакше – повернення до кроку 2.

7. Розрахувати цільову функцію помилки мережі, мінімізується:

$$E(w) = \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{P}(y_{ij}^{(N)} - d_{ij})^2 , \tag{5}$$

де y_{ij} – розрахований вихідний стан нейрона j вихідного шару N нейронної мережі при подачі на її входи *i-го* образу; d_{ij} – бажаний вихідний стан цього нейрона.

Якщо помилка мережі не перевищує деякий припустимий рівень помилки ε, мережа навчена й може виконувати завдання прогнозування, у противному випадку – повернення до кроку 2 (подача на входи мережі вхідних образів починається спочатку) [4].

Для корекції вагових коефіцієнтів на підставі експериментальних досліджень обраний алгоритм Resilient Propagation (RPROP), що використовує знаки часток похідних й індивідуальну для кожної ваги величину кроку відновлення, що адаптується в процесі навчання:

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^{+}\Delta_{ij}^{(t-1)}, & \dfrac{\partial E^{(t)}}{\partial w_{ij}}\dfrac{\partial E^{(t-1)}}{\partial w_{ij}} > 0 \\ \eta^{-}\Delta_{ij}^{(t-1)}, & \dfrac{\partial E^{(t)}}{\partial w_{ij}}\dfrac{\partial E^{(t-1)}}{\partial w_{ij}} < 0 \\ \Delta_{ij}^{(t-1)}, & \dfrac{\partial E^{(t)}}{\partial w_{ij}}\dfrac{\partial E^{(t-1)}}{\partial w_{ij}} = 0 \end{cases}, \quad \text{де} \quad 0 < \eta^{-} < 1 < \eta^{+} . \tag{6}$$

Зміна ваги провадиться за формулою:

$$\Delta w_{ij}^{t} = -\Delta_{ij}^{t}\,\mathrm{sign}\left(\frac{\partial E^{(t)}}{\partial w_{ij}}\right). \tag{7}$$

Початкова ініціалізація ваг виконується відповідно до методу Нгуєна-Відроу, що дозволяє значно скоротити кількість ітерацій навчання за рахунок більшої близькості початкових значень до оптимального [5].

Вибір структури нейронної мережі заснований на теоремі Колмогорова, відповідно до якої для моделювання будь-якого процесу досить нейронної мережі з одним схованим шаром з $2n+1$ нейронами в цьому шарі, однак передбачається дослідити з погляду підвищення якості рішення завдання доцільність використання нейроеволюційного алгоритму.

Весь вхідний набір даних розділяється на 2 непересічні множини. Одна з них – навчальна вибірка – використовується для навчання нейронної мережі, при цьому застосовується метод вікон, друга – контрольна вибірка – для перевірки якості прогнозу, що здійснюється за схемою, зображеною на рис. 3. Виявлення адекватної математичної моделі дозволить здійснити якісний прогноз поведінки залишків на клієнтських рахунках. Експерименти, що проведені під час досліджень, були спрямовані саме на виявлення параметрів такої моделі, здатної якомога краще прогнозувати суми залишків на рахунках клієнтів в умовах впливу вищевказаної множини факторів.

Для запобігання помилок у прийнятті рішень і економічних втрат необхідно періодично корегувати модель, заново навчаючи мережу з урахуванням нових практичних даних.

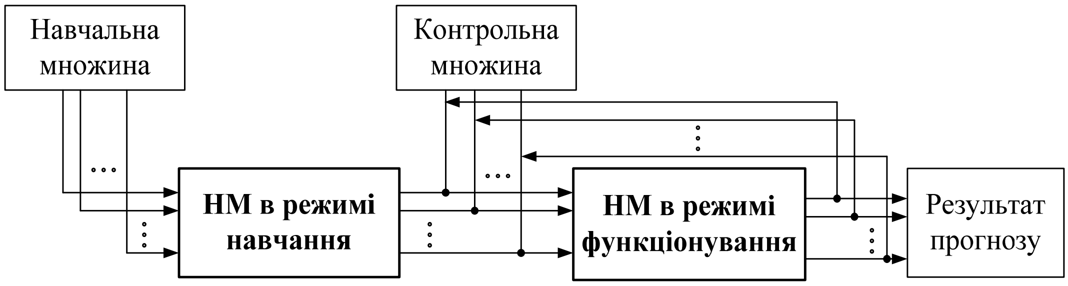


Рисунок 3 – Послідовність використання НМ для завдань прогнозування

На основі розглянутої моделі НМ зі зворотним поширенням похибки за допомогою розробленої комп'ютерної системи прогнозування був проведений ряд досліджень впливу структури НМ на швидкість її навчання й похибку прогнозування. При цьому ставилися такі задачі:

– пошук оптимального зображення й набору накоплених даних;

– пошук оптимального розміру вікна;

– пошук оптимальної структури мережі.

Метою експериментів було виявлення таких умов, при яких прогноз буде найбільш надійним.

Прогнозування здійснювалося на основі часової послідовності щомісячних середньозважених залишків коштів на депозитних рахунках у гривні й факторів, що впливають, з березня 2001 року по травень 2006 (63 виміри). Прогноз будувався для останніх шести місяців, реальні значення яких уже відомі.

При проведенні досліджень нормування вхідних значень здійснювалося в інтервалі [–0,5, 0,5], використовувалася сигмоїдальна функція активації нейронів зі зсувом. У ході попередніх експериментів зроблений висновок, що заміна функції активації на гіперболічний тангенс і інтервалу входів на [–1,1] не відбивається на якості навчання й прогнозних здатностей мережі, тому може також бути рекомендована для побудови моделі. Симетричний інтервал нормування зберігає рівноправність значень входів у процесі навчання.

Найкращі прогнозні властивості показали тришарові НМ із числом нейронів у схованому шарі від 7 до 40 і розміром вхідного вікна 5, а також чотиришарові, що містять до 20 нейронів у першому схованому шарі, і десь до 20 – у другому.

Встановлення ширини вікна у значення 7 та більше призводило до зниження передбачуваності, тобто збільшення розмірності входів вже не компенсувалося збільшенням їх інформативності. Оптимальна ширина вхідного вікна визначена на рівні 5.

Результати проведених експериментів свідчать про доцільність введення в модель додаткового вхідного фактора часу. У цілому НМ, побудована за останнім способом, найточніше прогнозує значення залишків коштів. Експерименти свідчать також про важливість попередньої обробки даних (підвищення ентропії входів привело до поліпшення узагальнюючих здатностей мережі). Незважаючи на те, що правильний підбір вхідних факторів має чимале значення, використання надмірності у вхідних даних привело в цілому до непоганих результатів.

На рис. 4 наведені результати ретроспективного прогнозування залишків коштів на депозитних рахунках клієнтів у гривні. Розмір вхідного вікна – 6, фактори, що враховуються, наведені в табл. 2, крім того, в модель включені попередні значення залишків, фактор часу. Структура нейронної мережі: 66 – 24 – 9 – 1.

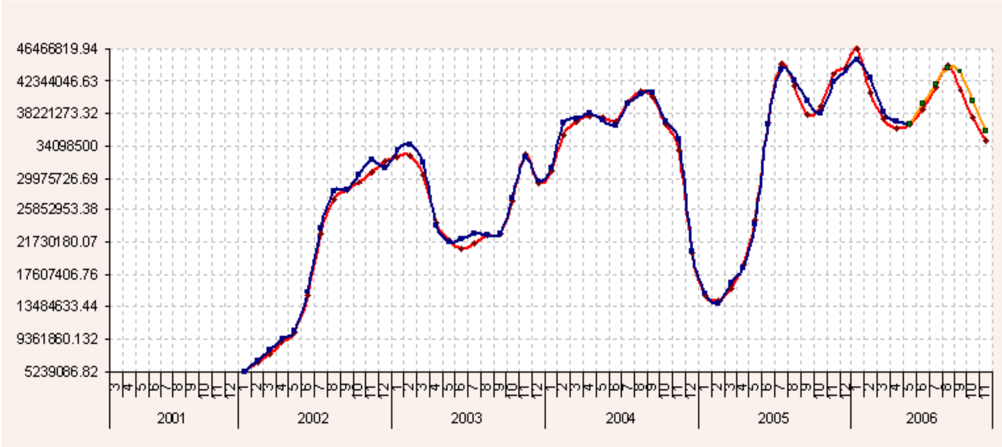


Рисунок 4 – Результати ретроспективного прогнозування залишків коштів на депозитних рахунках клієнтів у гривнях

Як видно, похибка передбачення в перший місяць становить менш 1%, що є дуже хорошим результатом. Незважаючи на те, що модельні значення не повністю збігаються із практичними, можна стверджувати, що тренд зміни шуканих величин відбитий точно, отже, використання даних моделей при прийнятті рішень цілком може себе виправдати. Єдина проблема, що встає перед виконавцем прогнозу – це проблема зваженого й обґрунтованого прогнозу макроекономічних показників країни.

Таблиця 2 – Результати прогнозування залишків коштів

Валюта	Гривня			Долар США		
	Модель	Прогноз	Відносна похибка, %	Модель	Прогноз	Відносна похибка, %
Червень	4226193	4253729	0,6515	38739749	39467827	1,8794
Липень	4647705	4614175	0,7214	41459900	41765493	0,9371
Серпень	5818321	5735748	1,4192	44468579	43987706	1,0814
Вересень	5382822	5373181	0,1791	41174561	43459369	5,5491
Жовтень	5073100	5235875	3,2086	37678554	39800746	5,6324
Листопад	4938519	5160394	4,4927	34785210	35899374	3,203

Таким чином експериментальним шляхом доведена ефективність пропонованого підходу до прогнозування залишків коштів на рахунках. Для підвищення якості прогнозу необхідно доповнити аналізовану інформацію додатковими даними. Крім того, у міру появи нових спостережень їхнє число зросте, що дозволить розв’язати проблему обмеженої кількості накопичених даних.

Незважаючи на ряд можливо неврахованих факторів (наприклад таких, як політична обстановка й інші не виражені кількісно показники), необхідність подальших тестувань і досліджень, можна сказати, що на основі запропонованої моделі можливе створення реальної системи прогнозування залишків коштів на поточних і депозитних рахунках клієнтів банку навіть в умовах нестабільного стану сучасної економічної системи.

Висновок

Володіння більш досконалими методами аналізу й прогнозування залишків на депозитних і поточних рахунках клієнтів допоможе банку більш ефективно й раціонально розподіляти свої зусилля й кошти на їхнє залучення, а також планувати використання

цих коштів як стабільний ресурс для кредитування (у тому числі міжбанківських кредитів), вкладення в цінні папери й т.д., максимізувати банківський прибуток шляхом інцестування прирівняних до власних коштів у прибуткові активи без ризику ліквідності. Побудована модель прогнозування залишків коштів на поточних і депозитних рахунках клієнтів дозволяє більш ефективно управляти банківськими коштами.

Література

1. Олійник Д. Управління ліквідністю комерційного банку на основі оптимального використання залишків тимчасово вільних коштів на рахунках клієнтів / Д. Олійник // Банківська справа. – 2001. – № 2. – С. 31-34, 41.
2. Костіна Н. Прогнозування надходження готівкових коштів до установи комерційного банку / Н. Костіна, П. Черняховська // Банківська справа. – 2000. – № 1. – С. 17-24.
3. Сергєєва Л. Комплексний аналіз залишків на рахунках клієнтів комерційного банку / Л. Сергєєва, А. Позднякова // Банківська справа. – 2002. – № 3. – С. 46-49.
4. Ежов А.А. Нейрокомпьютинг и его применения в экономике и бизнесе / А.А. Ежов, С.А. Шумский. – М. : МИФИ, 1998. – 224 с. – (Серия «Учебники экономико-аналитического института МИФИ»/ [под ред. проф. В.В. Харитонова]).
5. Nguyen D. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights / D. Nguyen, B. Widrow // Proceedings of the International Joint Conference on Neural Networks. – 1990. – Vol. 3. – P. 21-26.

Literatura

1. Olijnik D. Bankivs'ka sprava. 2001. № 2. P. 31-34, 41.
2. Kostina N. Bankivs'ka sprava. 2000. № 1. P. 17-24.
3. Sergeeva L. Bankivs'ka sprava. 2002. № 3. P. 46-49.
4. Ezhov A.A. Moscow : MIFI. 1998. 224 p.
5. Nguyen D. Proceedings of the International Joint Conference on Neural Networks. 1990. Vol. 3. P. 21-26.

Ю.К. Орлов, Е.М. Кирьян

Прогнозирование остатков денежных средств на текущих и депозитных счетах клиентов

В статье рассматриваются проблемы построения модели прогнозирования остатков денежных средств на текущих и депозитных счетах клиентов. Предлагается построение системы прогнозирования с помощью нейронных сетей, что позволяет повысить качество прогноза.

Yu.K. Orlov, E.M. Kirjan

Forecasting of Demand Balance on Current and Depositary Accounts of Clients

The article considers the problems of model construction of forecasting of demand balance on current and depositary accounts of clients. Construction of system of forecasting by means of neural networks that allows to raise quality of the forecast is offered.

Стаття надійшла до редакції 08.02.2011.